Course One
Foundations of Data Science

Instructions

Use this PACE strategy document to record decisions and reflections as you work through this end-of-course project. You can use this document as a guide to consider your responses and reflections at different stages of the data analytical process. Additionally, the PACE strategy documents can be used as a resource when working on future projects.

Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

- ☐ Complete the PACE Strategy Document to plan your project while considering your audience members, teammates, key milestones, and overall project goal.
- ☐ Create a project proposal for the data team.

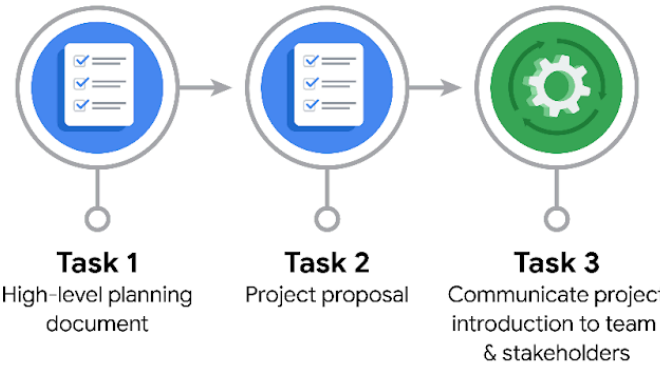
Relevant Interview Questions

Completing this end-of-course project will empower you to respond to the following interview topics:

- As a new member of a data analytics team, what steps could you take to get 'up to speed' with a current project? What steps would you take? Who would you like to meet with?
- How would you plan an analytics project?
- What steps would you take to translate a business question to an analytical solution?
- Why is actively managing data an important part of a data analytics team's responsibilities?
- What are some considerations you might need to be mindful of when reporting results?

Reference Guide

This project has three tasks; the following visual identifies how the stages of PACE are incorporated across those tasks.



Data Project Questions & Considerations

PACE: Plan Stage

- Who is your audience for this project?

The audience includes the Waze leadership team and data analytics team. The primary stakeholders are the decision-makers interested in reducing churn rates and improving user retention.

- What are you trying to solve or accomplish? And, what do you anticipate the impact of this work will be on the larger needs of the client?

You are building a machine learning model to predict user churn. The impact will help Waze take preventive measures to retain users, ultimately increasing user engagement and company growth.

- What questions need to be asked or answered?

  - What factors lead to user churn?
  - What data is available or missing to build an accurate model?
  - How will the model's performance be evaluated?
  - What strategies can be employed to reduce churn?

- What resources are required to complete this project?

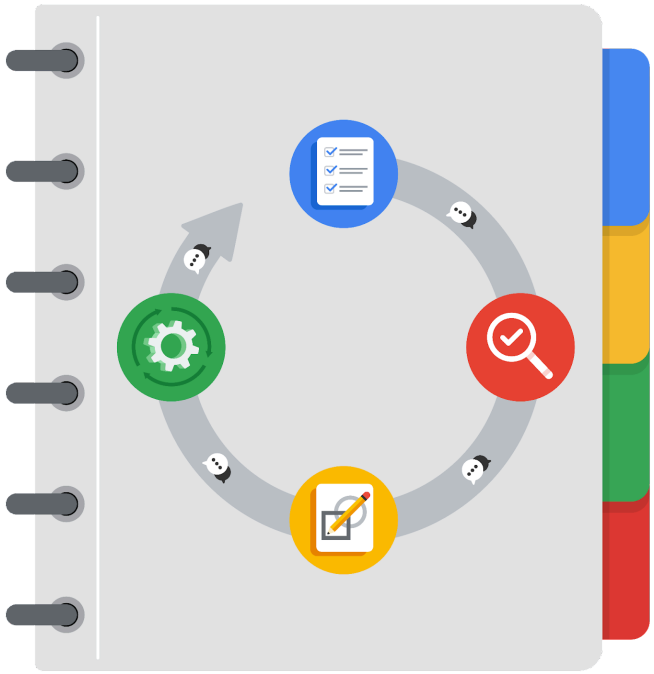
1. Waze user data (both structured and unstructured)

2. Analytical tools such as Python, R, or machine learning platforms (e.g., TensorFlow, Scikit-learn)
3. Team expertise in data analysis, machine learning, and communication of insights.

- What are the deliverables that will need to be created over the course of this project?

  - A machine learning model for churn prediction
  - A summary report of the data exploration, cleaning, and modeling process
  - Visualizations to illustrate key insights
  - A final presentation of findings and recommendations to stakeholders.

THE PACE WORKFLOW



[Alt-text: The PACE Workflow with the four stages in a circle: plan, analyze, construct, and execute.]

You have been asked to demonstrate for the company's data team how you would use the PACE workflow to organize and classify tasks for the upcoming project. Select a PACE stage from the dropdown buttons. A few tasks involve more than one stage of the PACE workflow. Additionally, not every workplace scenario will require every task. Refer back to the Course 1 end-of-course portfolio project overview reading if you need more information about the tasks within the project.

Project tasks

Following are a group of tasks your company's data team has determined need to be completed within this project. The data analysis manager has asked you to organize these tasks in preparation for the project proposal document. First, identify which stage of the PACE workflow each task would best fit under using the

drop down menu. Next, give an explanation of why you selected the stage for each task. Review the following readings to help guide your selections and explanation: The PACE stages and Communicate objectives with a project proposal. You will later reorder these tasks within a project proposal.

1. **Evaluating the model:** Execute

   Why did you select this stage for this task?

Model evaluation occurs after the model is built and is an essential step to ensure accuracy and reliability.

2. **Conduct hypothesis testing:** Construct **and** Analyze

   Why did you select these stages for this task?

Hypothesis testing helps determine if patterns in the data are statistically significant, which is part of the analysis stage.

3. **Begin exploring the data:** Plan

   Why did you select this stage for this task?

Initial data exploration is part of the planning phase, where you start understanding the structure and contents of the data.

4. **Data exploration and cleaning:** Analyze **and** Construct

   Why did you select these stages for this task?

Data cleaning is crucial to both analyzing and constructing accurate models since clean data is essential for robust analysis.

5. **Establish structure for project workflow (PACE):** Plan

   Why did you select this stage for this task?

Setting up the project structure and workflow falls under the planning phase as it lays the foundation for the project's progress.

6. **Communicate final insights with stakeholders:** Execute

   Why did you select this stage for this task?

Communicating findings happens toward the end of the project when results are available, making this part of the execution phase.

7. **Compute descriptive statistics:** Analyze

   Why did you select this stage for this task?

Descriptive statistics are part of the analysis phase as they summarize the key features of the data before building models.

8. **Visualization building:** Construct **and** Execute

   Why did you select these stages for this task?

Building visualizations supports both constructing the model and executing the final presentation of results.

9. **Write a project proposal:** Plan

   Why did you select this stage for this task?

Writing the proposal outlines the project's scope, objectives, and workflow, making it part of the planning phase..

10. **Build a regression model:** Construct **and** Analyze

    Why did you select this stage for this task?

Constructing a regression model requires both analyzing the data and building the actual machine learning model.

11. **Compile summary information about the data:** Analyze

    Why did you select this stage for this task?

Compiling data summaries is part of analyzing and understanding the dataset before building models.

12. **Build machine learning model:** Execute

Why did you select this stage for this task?

Building the model itself falls under the construct phase, where you create the predictive tool.